Evidence of the SM Higgs Boson in CMS in the Decay Channel into Tau Leptons

JHEP 05 (2014) 104

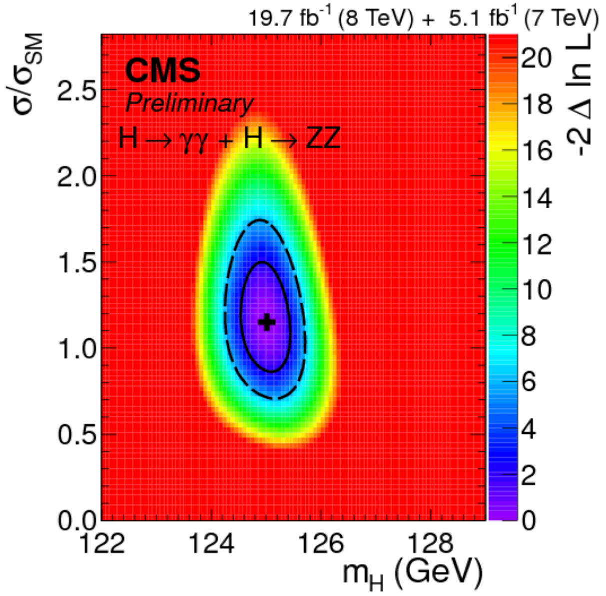
Higgs Hunting Workshop
July 21, 2014

Armin Burgmeier (DESY)
for the CMS collaboration

Karlsruhe Institute of Technology

Motivation

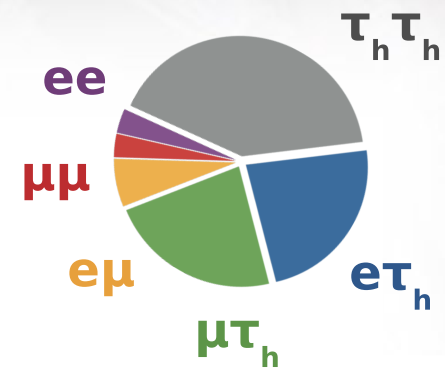
- A **Higgs boson has been found** at a mass of ~125 GeV
  - Signals have been seen in
    $H \to \gamma\gamma$ (5.6σ),
    $H \to ZZ$ (6.5σ),
    $H \to WW$ (4.7σ)
- Coupling to **Fermions**?
  - Fundamentally different than coupling to bosons
  - Only indirect evidence from bosonic channels
  - **Down-type** fermion couplings can be probed with $H \to \tau\tau$



CMS-PAS-HIG-14-009

Overview

- High **backgrounds**, dominated by:



Z → ττ:

Normalized to measured cross section, **Shape** from embedded data events

Z → ee/μμ:

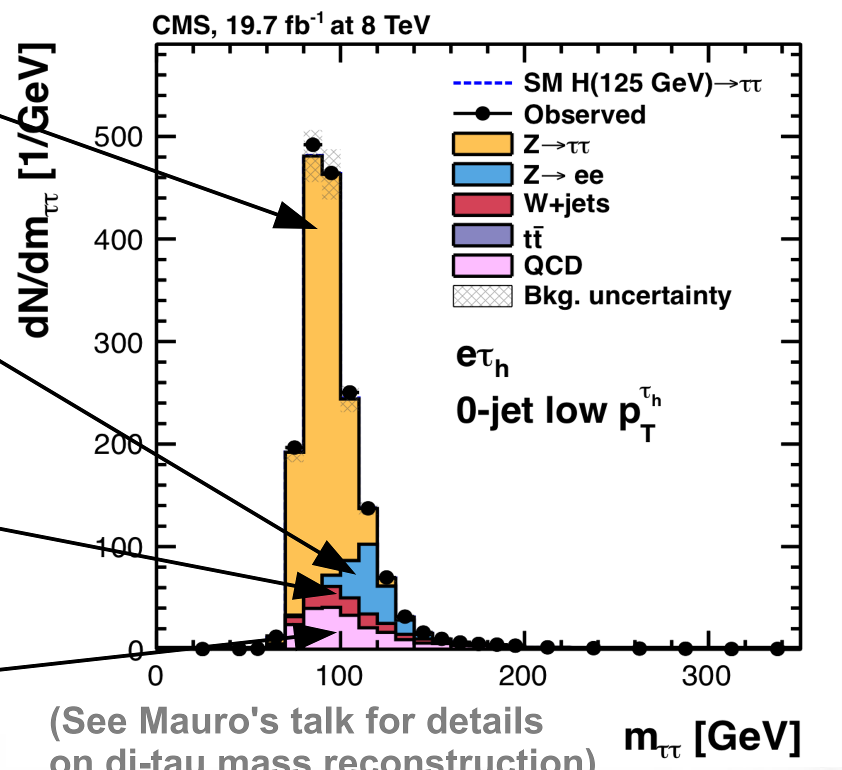
Shape and **Normalization** from fit to data in 0-jet category

W + Jets:

Normalized to data in high m_T sideband, **Shape** from Simulation

QCD Multijet:

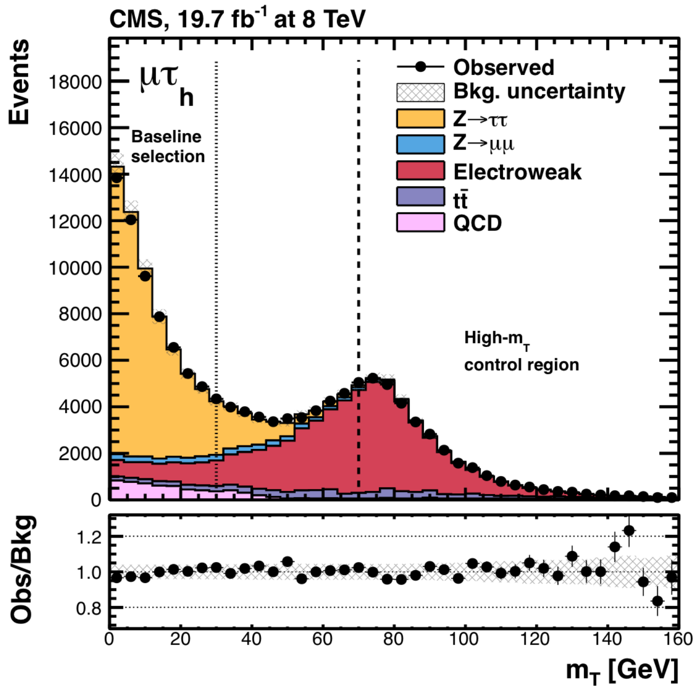
Shape and **Normalization** from Same Sign data



(See Mauro's talk for details on di-tau mass reconstruction)

Background Estimation

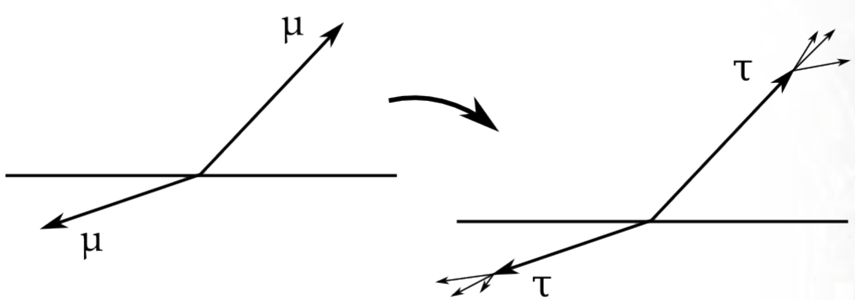
W + Jets:

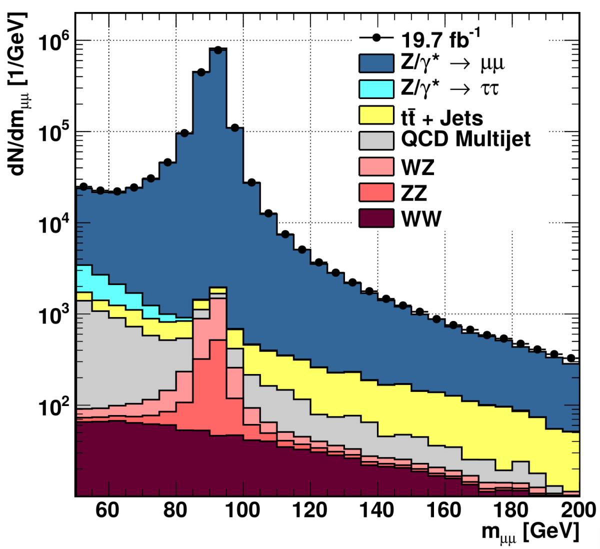


$$m_{\mathrm{T}} = \sqrt{2 p_{\mathrm{T}} E_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos\Delta\phi)}$$

pure W + Jets sideband

Z → ττ:





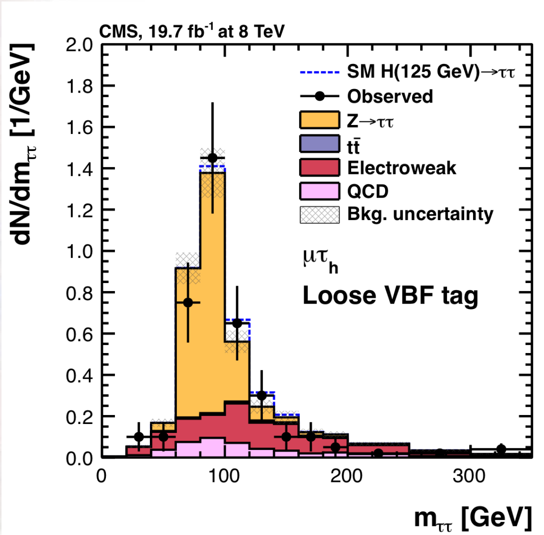
very pure di-muon selection

IEKP-KA/2014-09

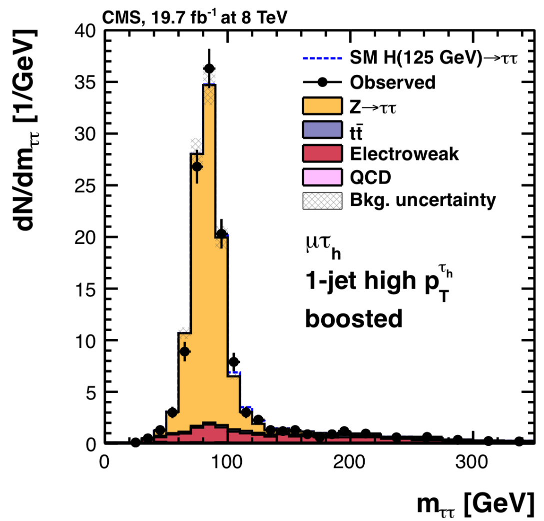
Event Categorization

VBF: | **1 Jet:** | **0 Jets:**

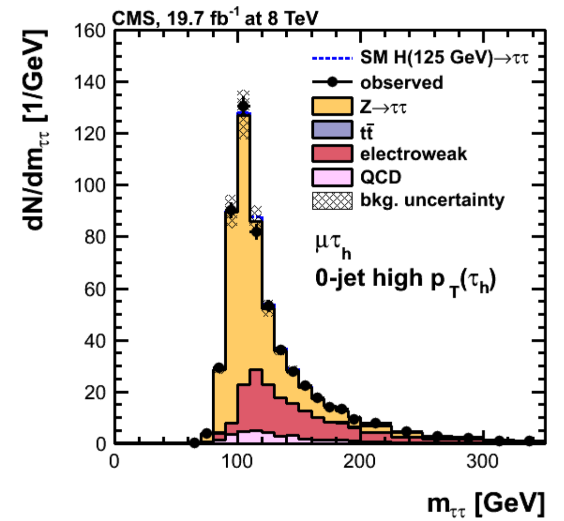
JHEP 05 (2014) 104



- Low event statistics
- High S/B



- Exploit boost of the Higgs system: Improved mass resolution

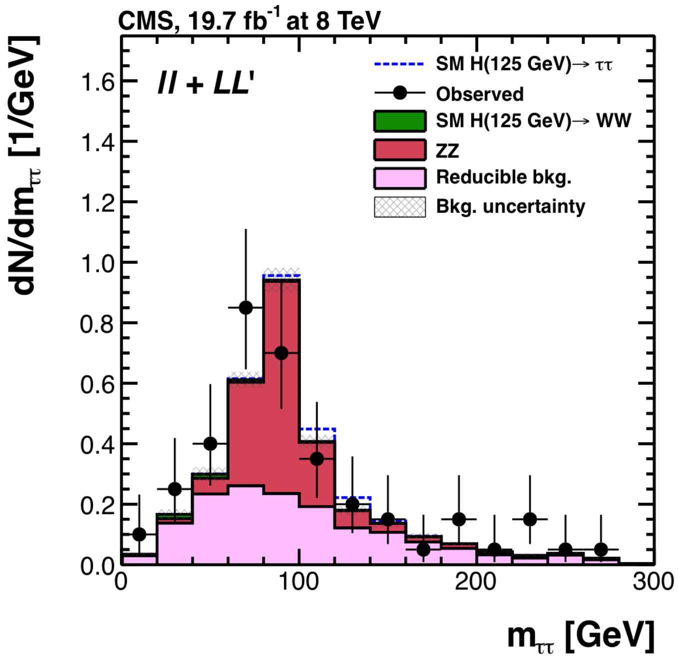


- Low S/B
- Important for Constraining Nuisance Parameters

$\mu\tau_h$ channel

VH Associated Production

- More than 2 leptons in the event
- Easy to **trigger**
- Low SM Background
- But: Low cross section
- WZ/ZZ is **irreducible background**
  - From **simulation**
- Other background have **misidentified leptons**
  - estimated from **data**



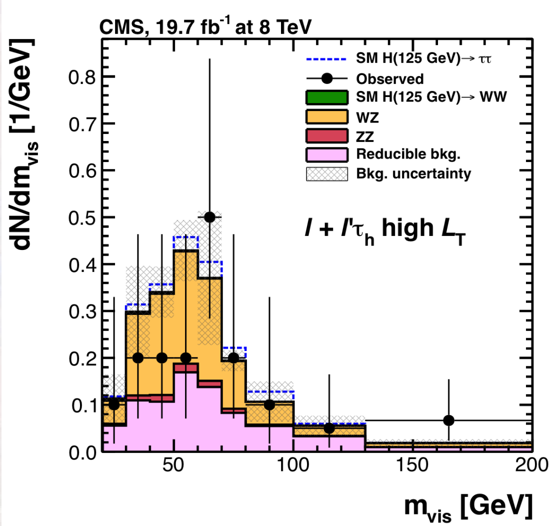
Channels analyzed:

$Z(ee, \mu\mu)H(\mu\tau_h, e\tau_h, \tau_h\tau_h, e\mu)$

$W(e\nu, \mu\nu)H(\mu\tau_h, e\tau_h, \tau_h\tau_h)$

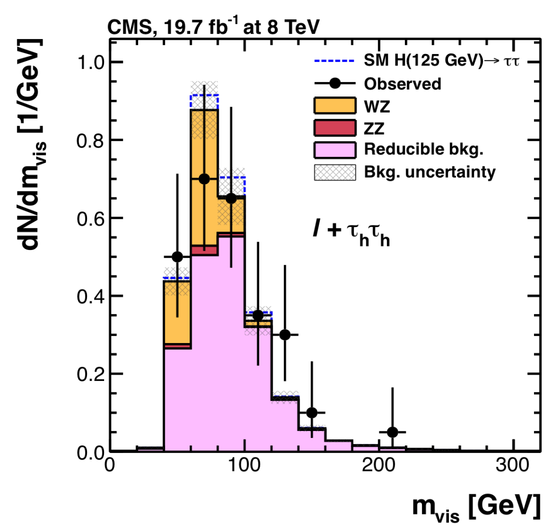
VH Categorization

WH semi-leptonic:



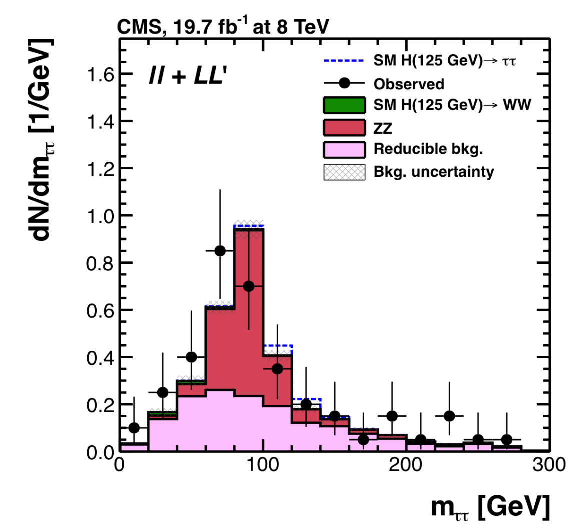
- 3 Leptons
- Dominated by WZ pair production

WH hadronic:



- 3 Leptons
- Dominated by mis-identified jets

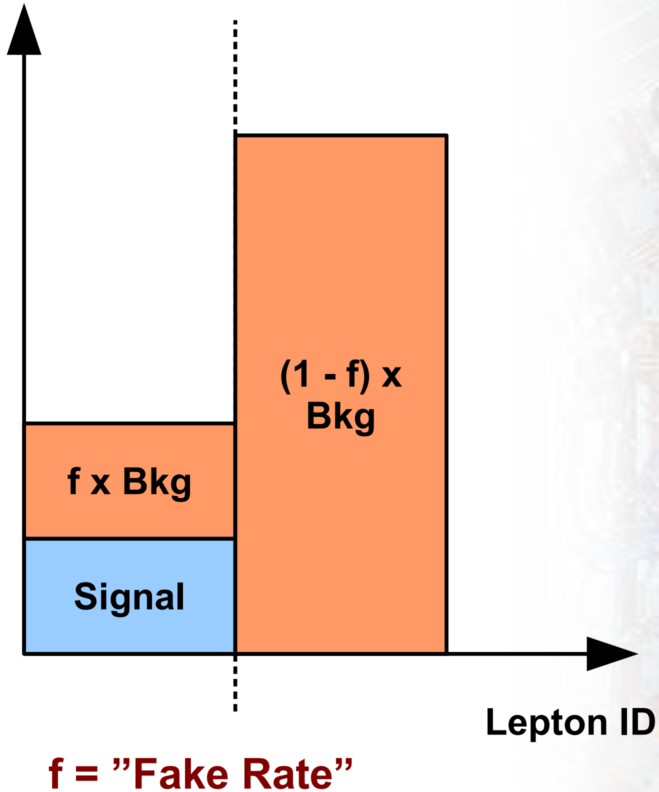
ZH:



- 4 Leptons

VH Background Estimation

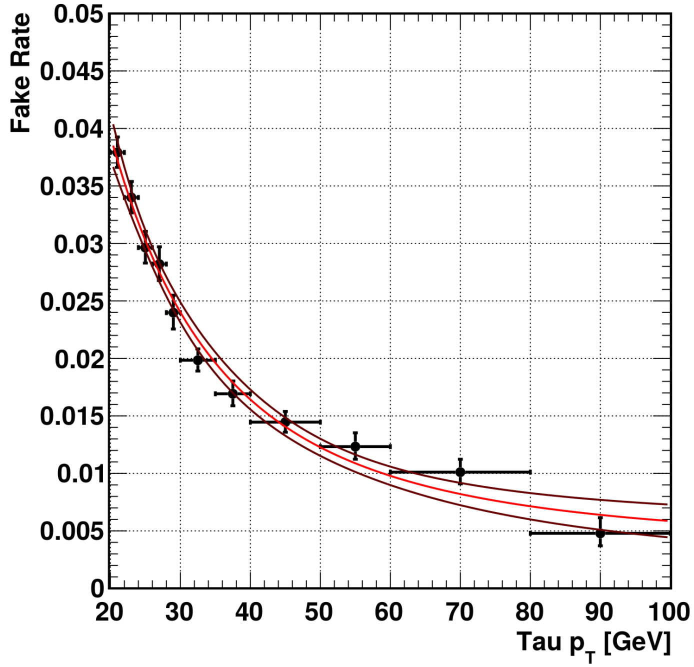
- Major background from **misidentified leptons**
- Select background-enriched region by **inverting lepton ID** or isolation
- Weighting with ”Fake Rate” gives estimation in signal region
- Works also in 2D for backgrounds with different misidentified leptons
  - e.g. e+μ in $e\mu\tau_h$ channel



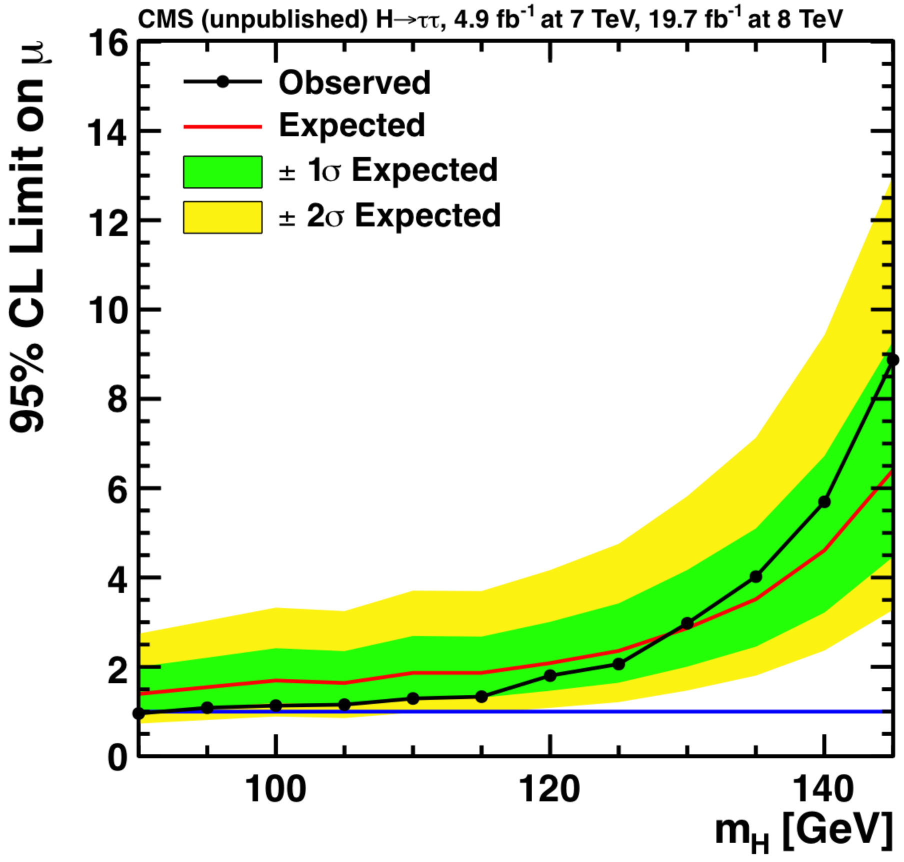
Fake Rate Measurements

- Fake Rate measured in well-known **control regions**
  - For example, $Z \rightarrow \mu\mu$ events
- Fake Rate depends on **many parameters**
  - p_T and η of the lepton
  - Jet multiplicity
  - Hard physics process
- Major source of **systematic uncertainty**
  - 20-30% in most channels

IEKP-KA/2014-09

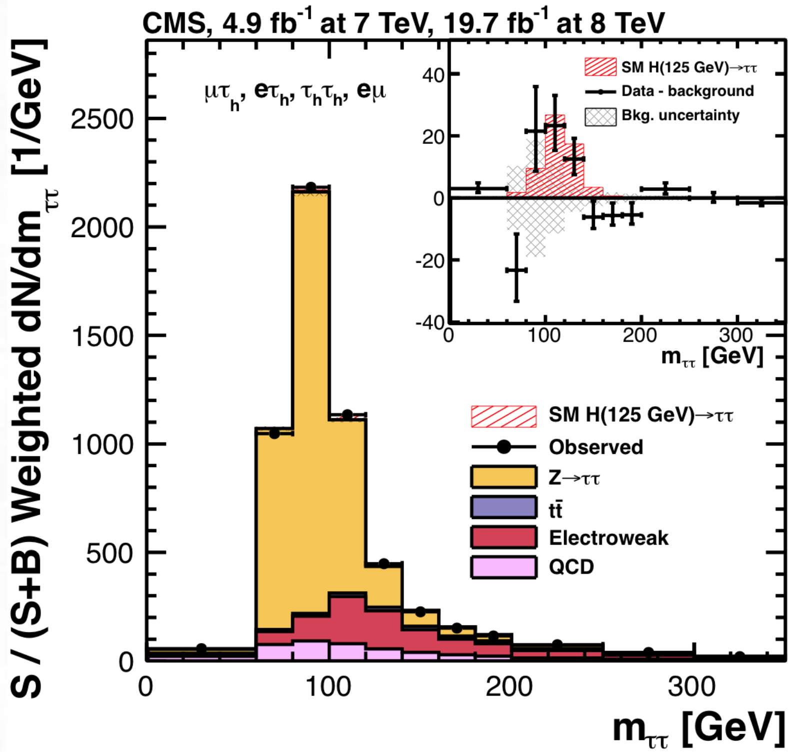


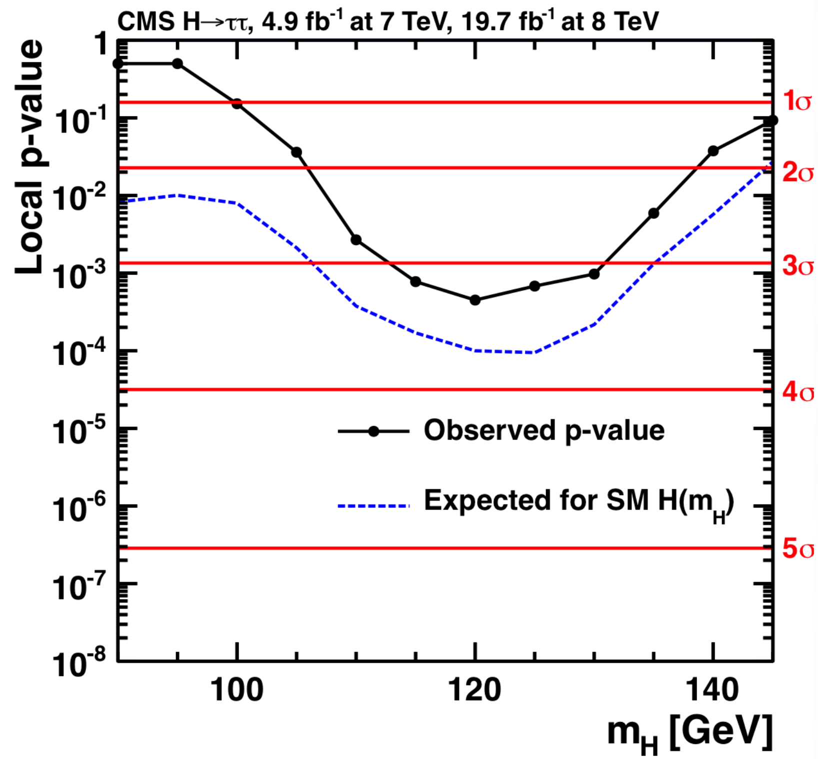
VH Result



- No signal observed in VH channels
- Sensitivity dominated by **statistical precision**
- High sensitivity at **low mass**

Full Combination

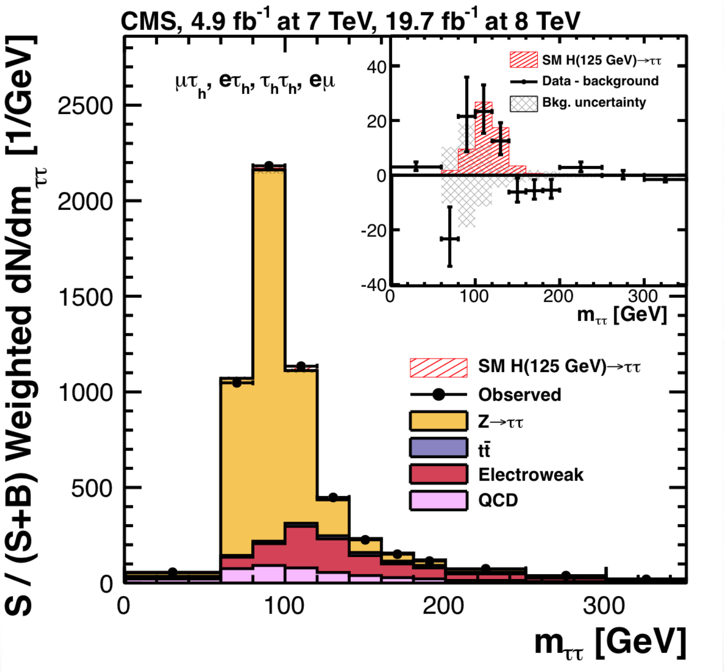




- Signal builds up slowly in various **channels and categories**
  - Visualized in **S/(S+B) weighted** mass distribution
- CMS sees **3.2σ** evidence for the H → ττ decay

Conclusions

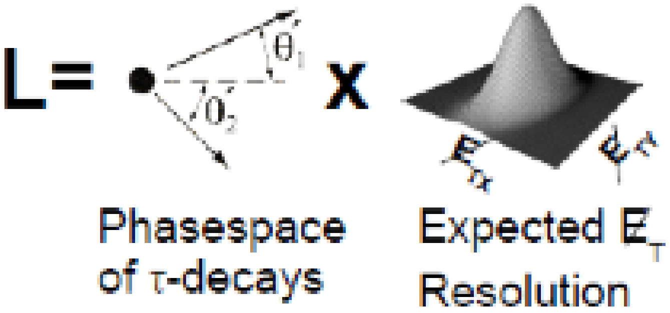
- **SM Higgs Results** in the ττ channel have been presented
  - Analysis is complex due to high backgrounds and the **combination of many channels** and categories
  - Different analysis strategy in VH channels
- CMS sees an **excess around 125 GeV** at **3.2σ** significance!
- Evidence for coupling of Higgs boson to **tau leptons**!
- More on H → ττ and other fermionic channels in **Mauro's talk**



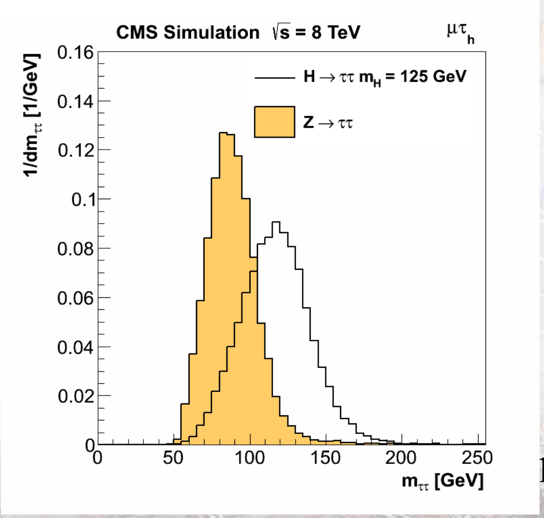
Backup

Di-tau mass reconstruction

- Use di-tau mass as **discriminating variable**
- Undetected neutrinos lead to **underestimation** of the di-τ mass
- Likelihood-based method to find mass which is **most compatible** with:
  - Tau decay kinematics
  - Visible decay products
  - E_T^{miss} + uncertainty

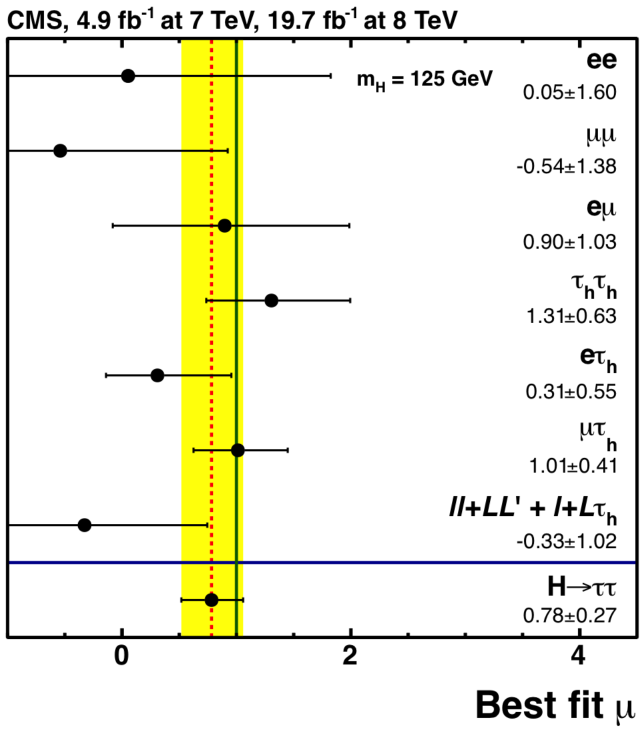


- Mass resolution:
  - 10% to 20% (depending on final state)

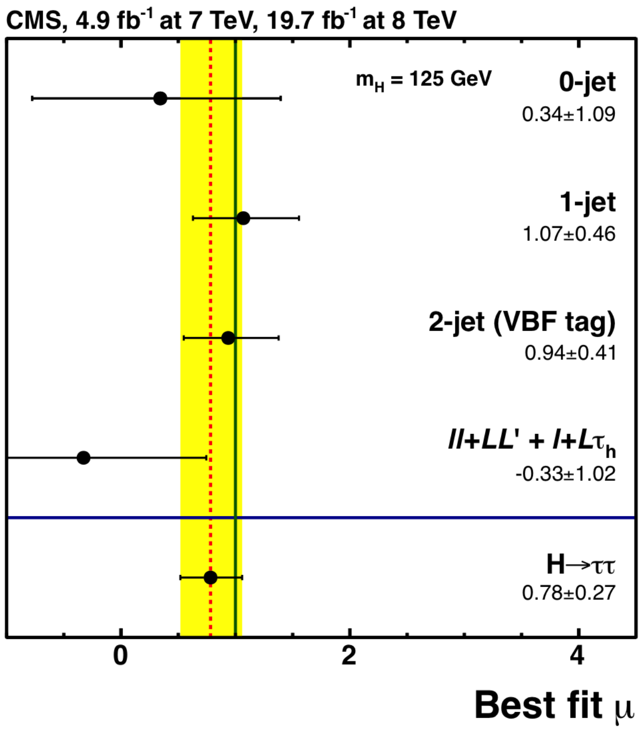


Best Fit Signal Strength

By channel:



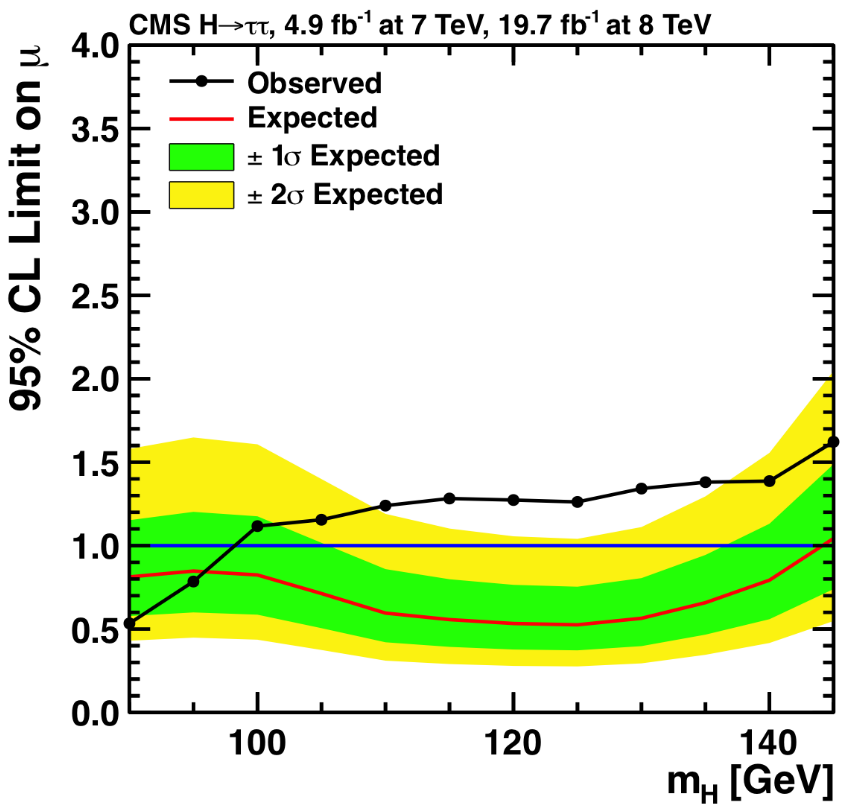
By category:



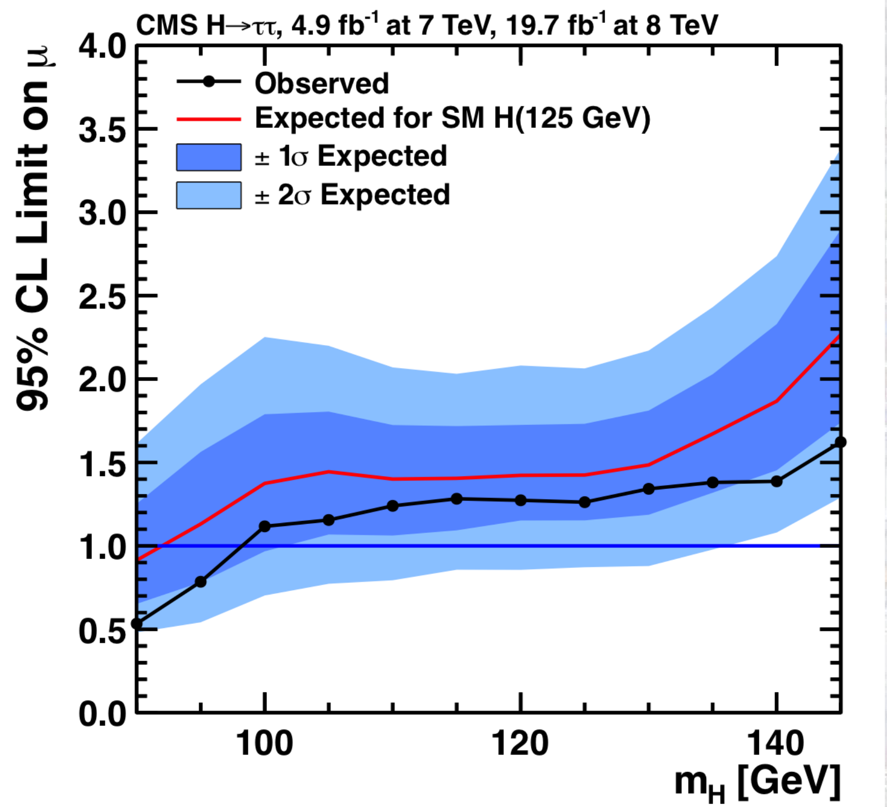
- Important nuisance parameters **shared** between channels and categories (**constrained** by high statistics categories in global fit)
- Best fit **μ = 0.79 ± 0.27**

Expected Exclusion Limits

Background only:

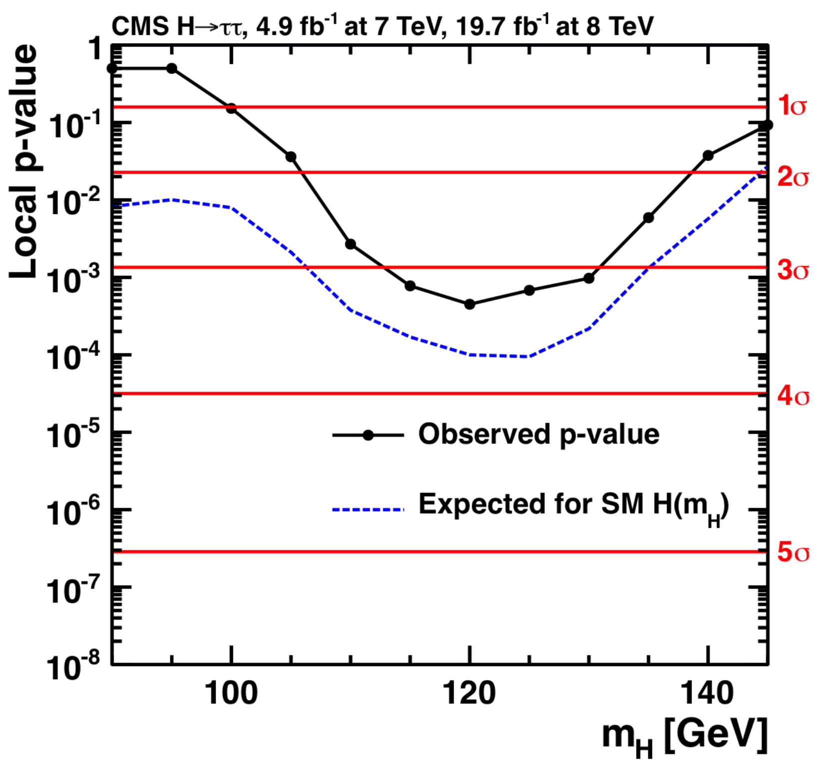


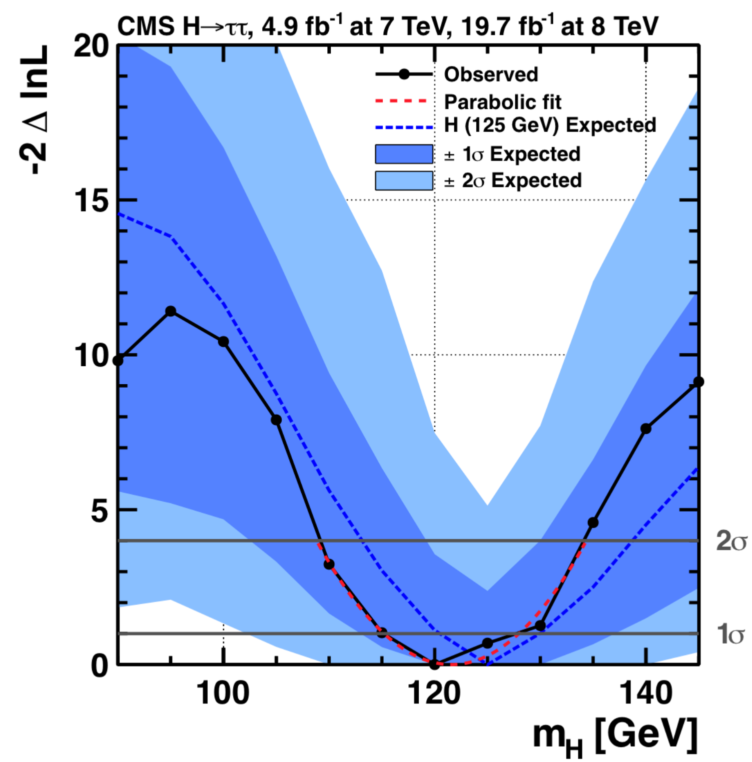
1x SM + Background:



- Excess is **compatible with SM Higgs** boson hypothesis over wide mass range

p-value and Mass Scan

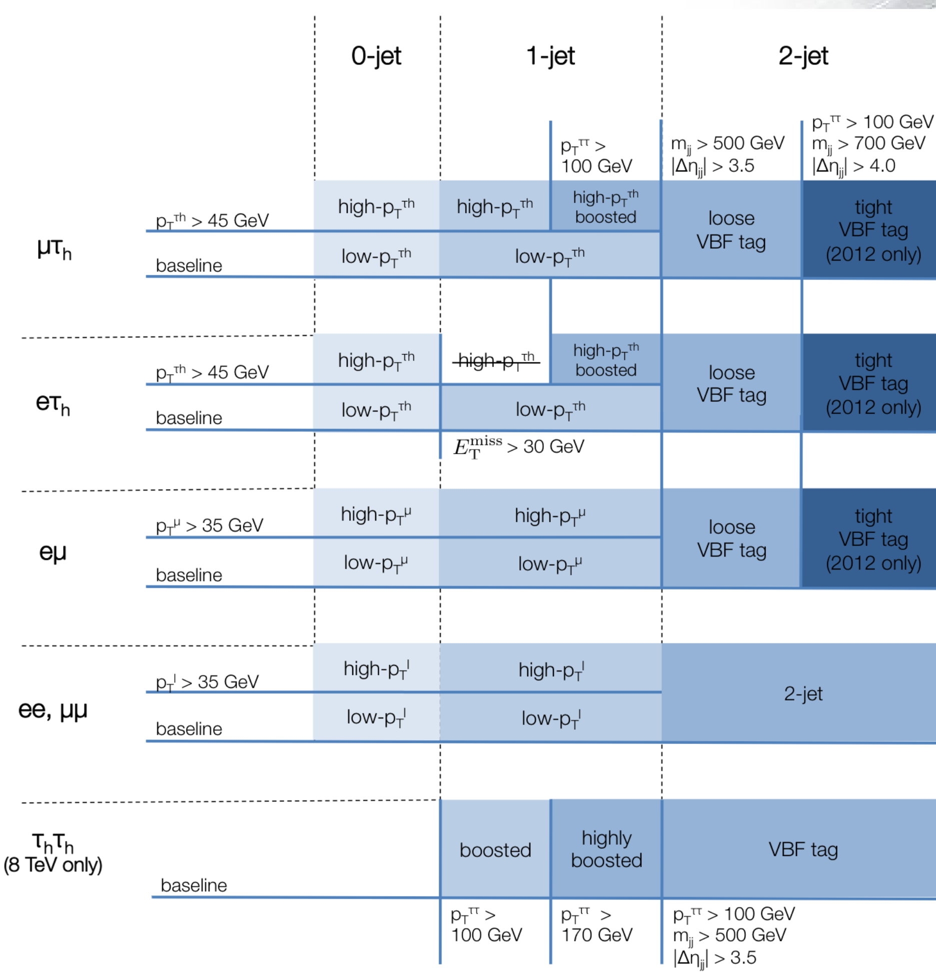




- Largest **observed significance** (3.3σ) at m_H = 120 GeV
- **Mass scan**: **m_H = 122 ± 7 GeV**

Event Categorization

- Use full event kinematics to **categorize events**, based on
  - jet multiplicity
  - $p_T^{\tau\tau} = |\vec{p}_T(L) + \vec{p}_T(L') + E_T^{miss}|$
  - $p_T(\tau_h / \tau_l)$
- Optimizes **overall sensitivity**
- Less categories in the 7 TeV data
- **58 categories** in total
  - Fit for signal in all of them



Categories in 8 TeV

Event Categorization at 7 TeV

Channel	Selection	0-jet	1-jet		2-jet
				$p_T^{\tau\tau} > 100$ GeV	$m_{jj} > 500$ GeV, $\lvert\Delta\eta_{jj}\rvert > 3.5$
$\mu\tau_h$	$p_T(\tau_h) > 45$ GeV	high $p_T(\tau_h)$	high $p_T(\tau_h)$	high $p_T(\tau_h)$ boost	VBF tag
	baseline	low $p_T(\tau_h)$	low $p_T(\tau_h)$		
$e\tau_h$	$p_T(\tau_h) > 45$ GeV	high $p_T(\tau_h)$	high $p_T(\tau_h)$		VBF tag
	baseline	low $p_T(\tau_h)$	low $p_T(\tau_h)$		
			$E_T^{miss} > 30$ GeV		
$e\mu$	$p_T(\mu) > 35$ GeV	high $p_T(\mu)$	high $p_T(\mu)$		VBF tag
	baseline	low $p_T(\mu)$	low $p_T(\mu)$		
ee, $\mu\mu$	$p_T(l) > 35$ GeV	high $p_T(l)$	high $p_T(l)$		2-jet
	baseline	low $p_T(l)$	low $p_T(l)$		

Analysis Strategy

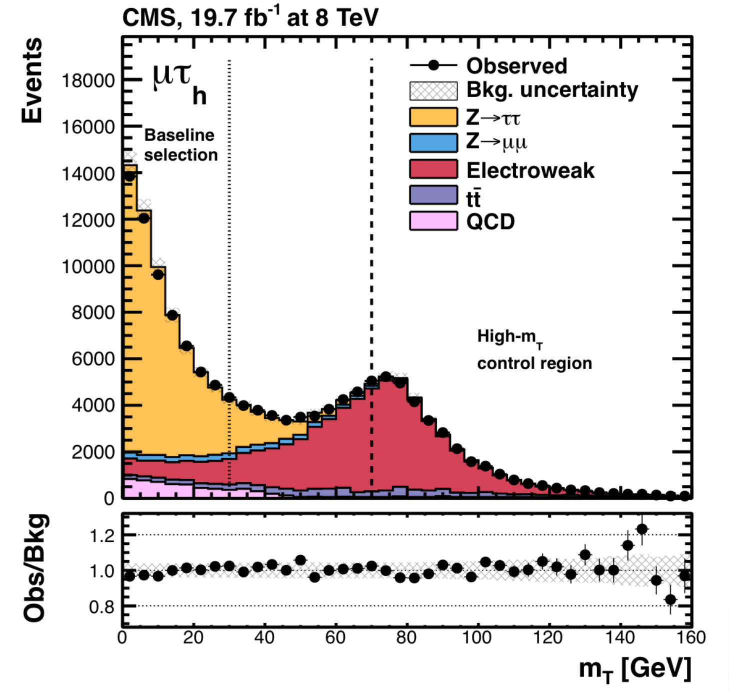
- Goal: Measure **coupling** of $H \rightarrow \tau\tau$ ($H \rightarrow WW$ is background)
- Final states with **2 leptons**: e, μ or τ_h (More than two leptons in VH)
- Light leptons from tau decays are **soft**
  - Need **low p_T thresholds** (→ cross triggers)

Channel	Offline p_T Threshold
$\mu\tau_h$	$p_T(\mu) > 20$ GeV, $p_T(\tau_h) > 30$ GeV
$e\tau_h$	$p_T(e) > 24$ GeV, $p_T(\tau_h) > 30$ GeV
$\tau_h\tau_h$	$p_T(\tau_h) > 45$ GeV
ee, eμ, μμ	$p_T(l_1) > 20$ GeV, $p_T(l_2) > 10$ GeV

- Isolated leptons to suppress e.g. QCD multijet events with jets misidentified as leptons
- $M_T(l, E_T^{miss}) < 30$ GeV to suppress W+Jets events

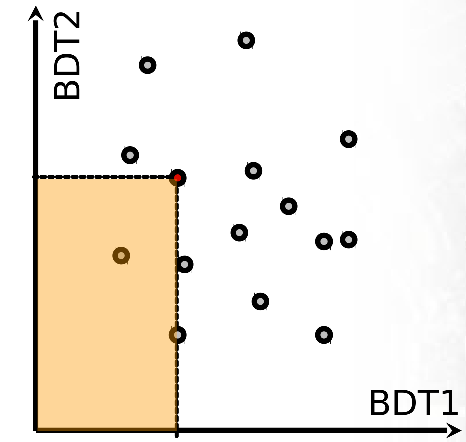
Background Rejection

- Very channel specific in general
  - Differentiate between
    - Irreducible backgrounds (same final state)
    - Reducible backgrounds (one or more objects misidentified)
- Main backgrounds:
  - $Z \to \tau\tau$
  - $Z \to ee/\mu\mu$
  - W + Jets
  - QCD Multijet
  - $t\bar{t}$

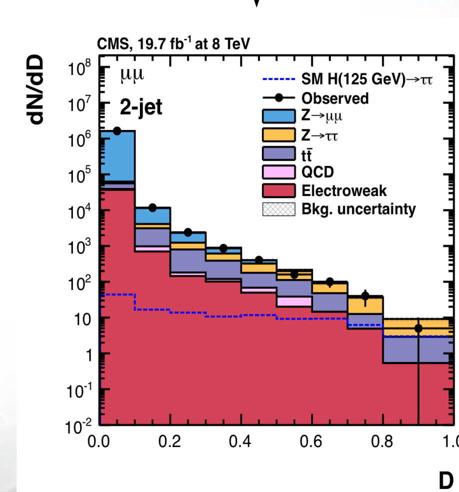


Same Flavor Dilepton Channels

- Different analysis strategy
- No τ_h reconstruction needed
- Additional **direct Z → ll** background
- Train **two BDTs**
  - BDT1: Separate Z → ll from Z/H → ττ
  - BDT2: Separate Z → ττ from H → ττ

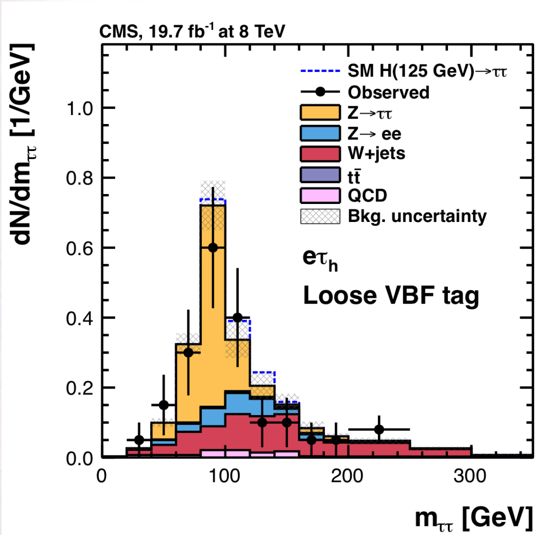


$$D_{\mathrm{cat}} = \int_0^{\mathrm{BDT}_1} \int_0^{\mathrm{BDT}_2} f_{\mathrm{cat}}^{\mathrm{sig}}(\mathrm{BDT}'_1, \mathrm{BDT}'_2)\, \mathrm{dBDT}'_1\, \mathrm{dBDT}'_2$$



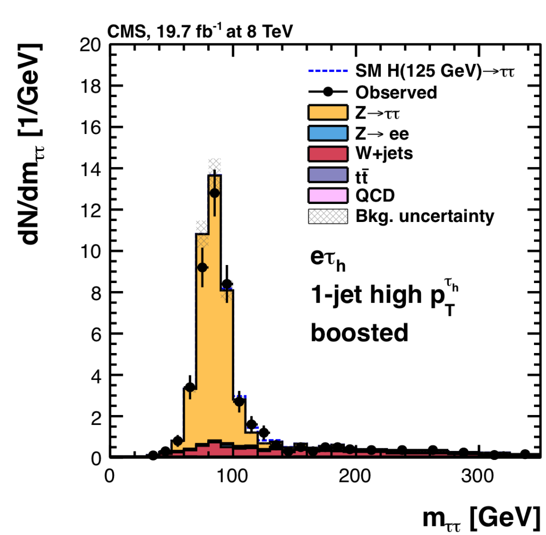
Di-Tau Mass Distributions ($e\tau_h$)

VBF:



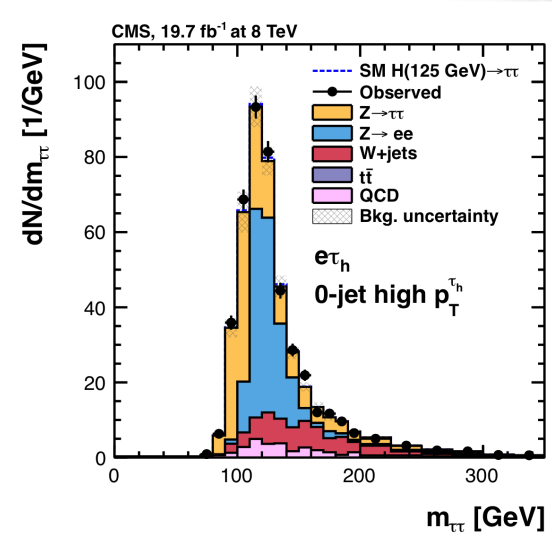
- Low event statistics
- High S/B

1 Jet:



- Exploit boost of the Higgs system: Improved mass resolution

0 Jets:

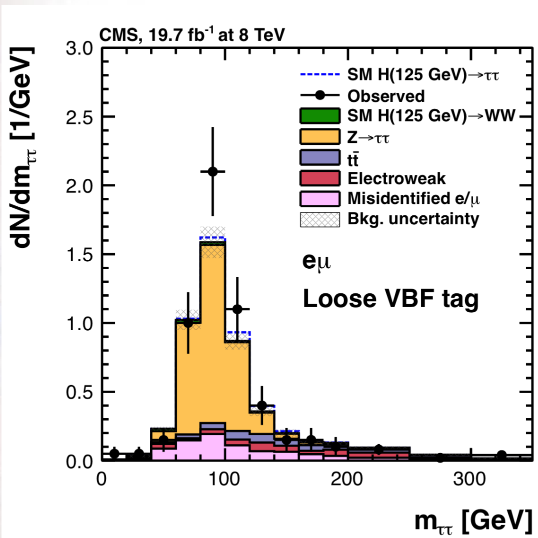


- Low S/B
- Important for Constraining Nuisance Parameters

$e\tau_h$ channel

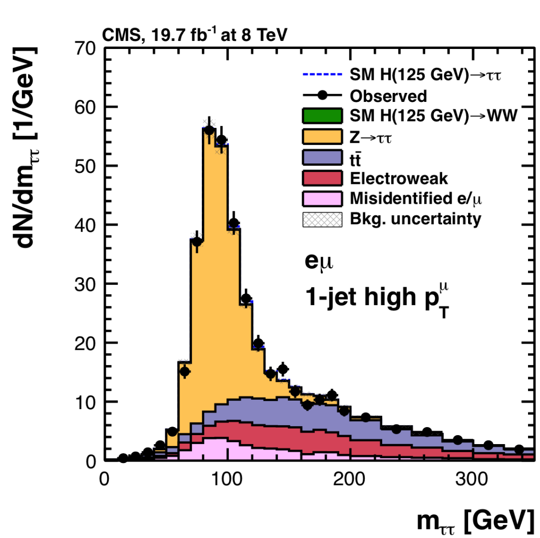
Di-Tau Mass Distributions (eμ)

VBF:



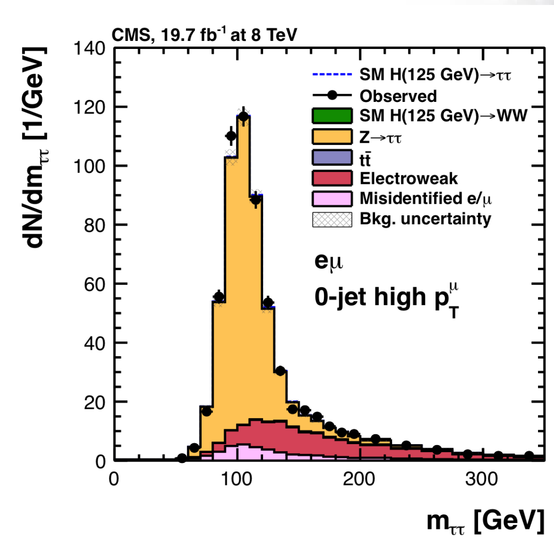
- Low event statistics
- High S/B

1 Jet:



- Exploit boost of the Higgs system: Improved mass resolution

0 Jets:

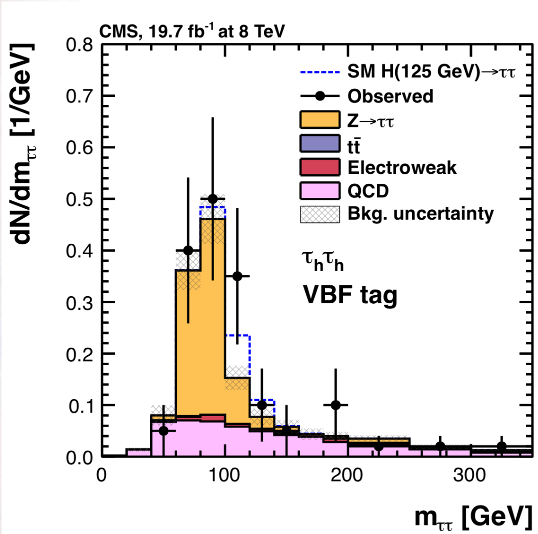


- Low S/B
- Important for Constraining Nuisance Parameters

eμ channel

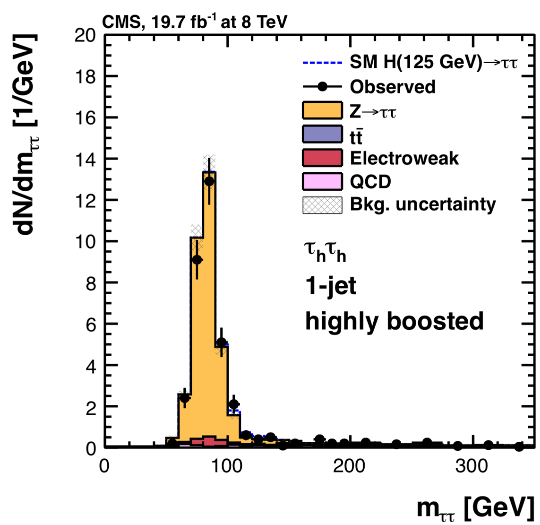
Di-Tau Mass Distributions ($\tau_h \tau_h$)

VBF:



- Low event statistics
- High S/B

1 Jet:



- Exploit boost of the Higgs system: Improved mass resolution

0 Jets:

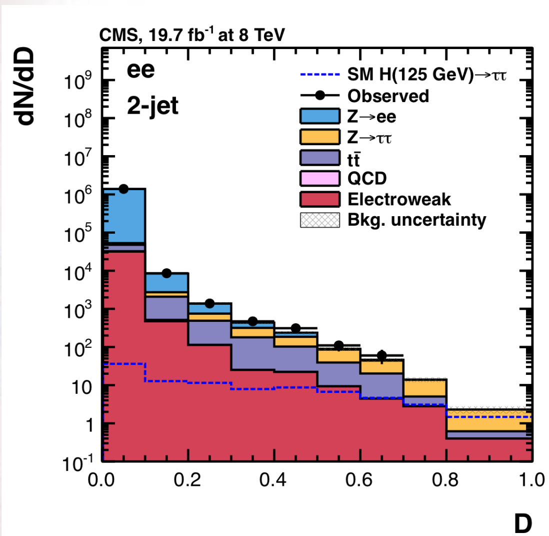
No 0-Jet category due to trigger requirements in this channel

- ~~Low S/B~~
- ~~Important for Constraining Nuisance Parameters~~

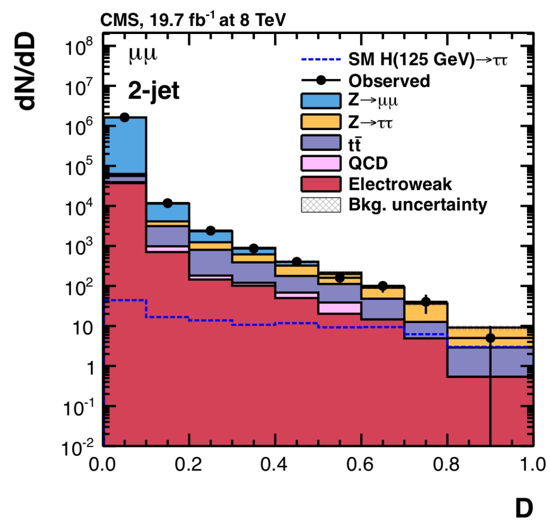
$\tau_h \tau_h$ channel

Comb. BDT Distributions (mm/ee)

ee VBF:

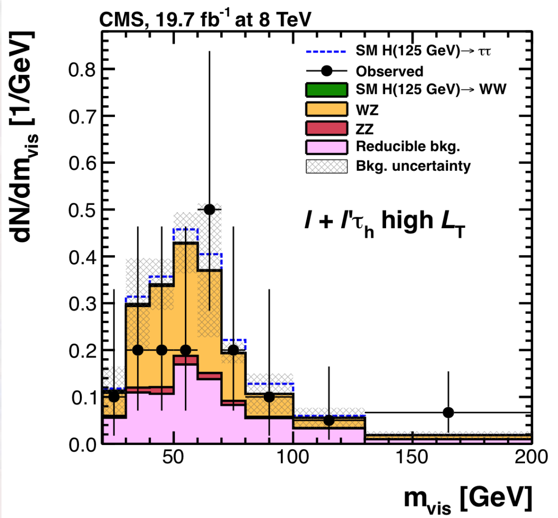


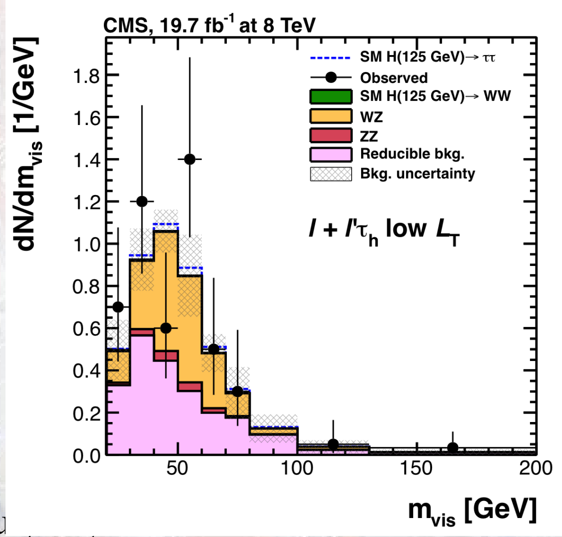
mm VBF:



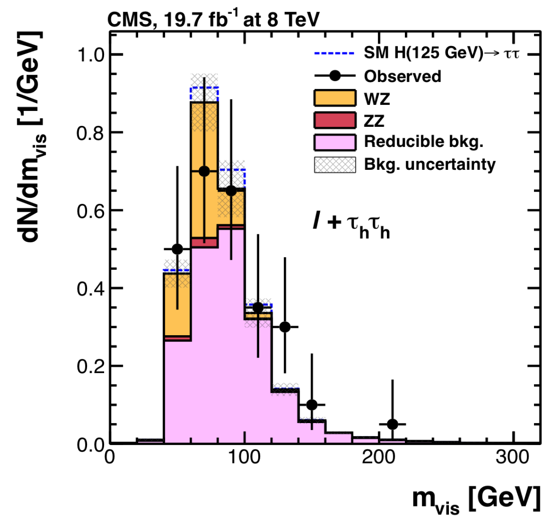
Di-Tau Mass Distributions (VH)

WH semi-leptonic:

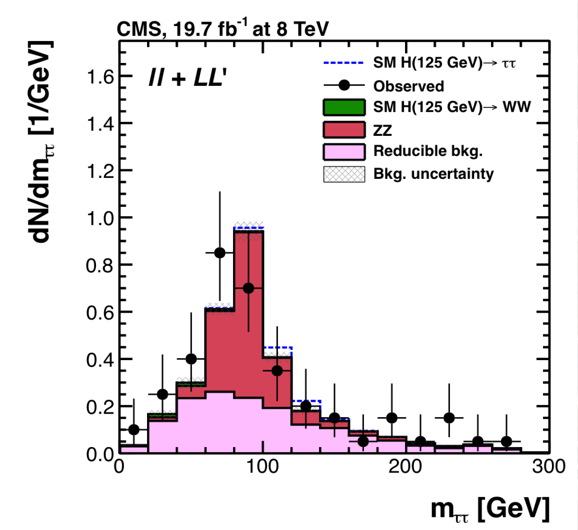




WH hadronic:



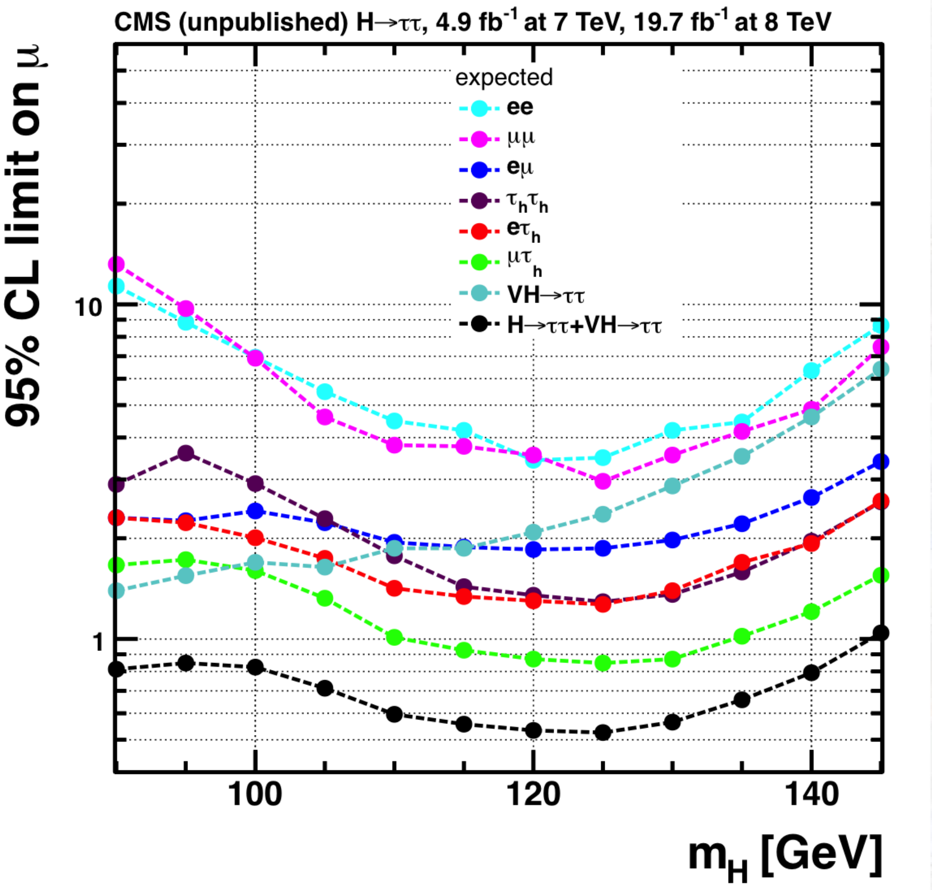
ZH:



VH channels

Expected Limit by Channel

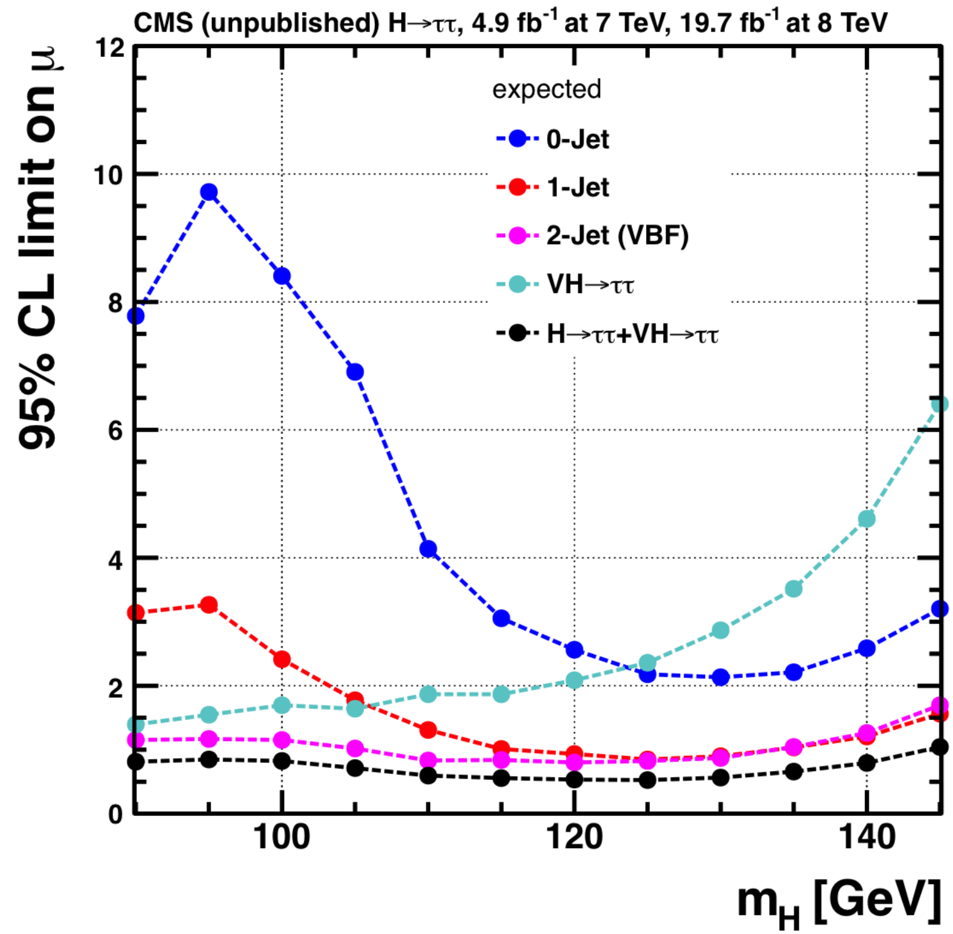
HIG-13-004



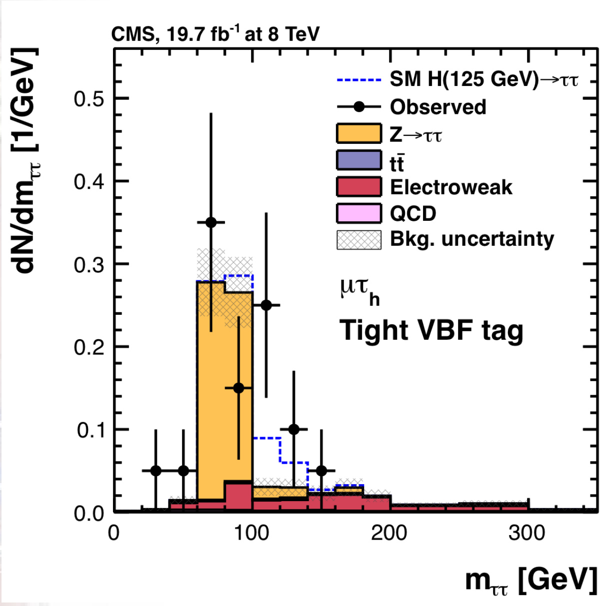
- **Combine all channels** and categories for statistical interpretation
- 95% C.L. Frequentist **Exclusion Limits** are set with the CLs method
- VH channels sensitive at low masses

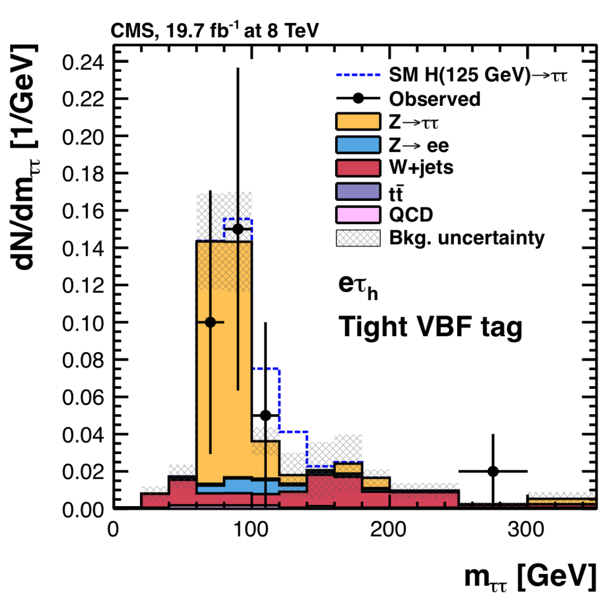
Only combination of channels is sensitive to SM Higgs

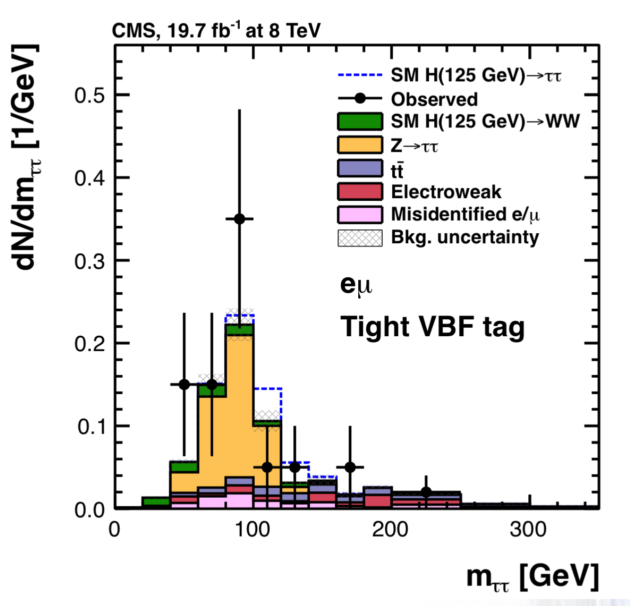
Expected Limit By Category



Tight VBF Category

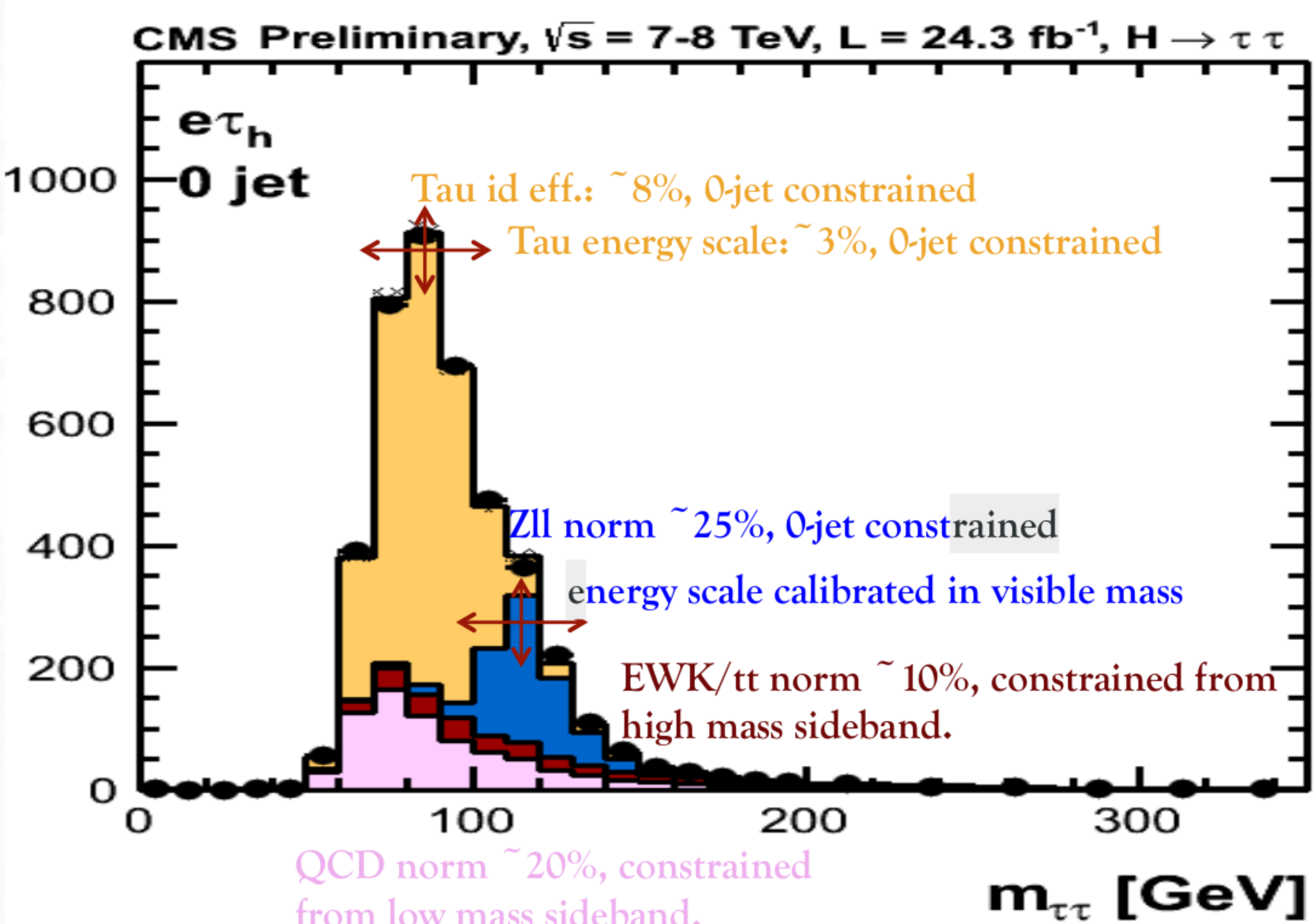






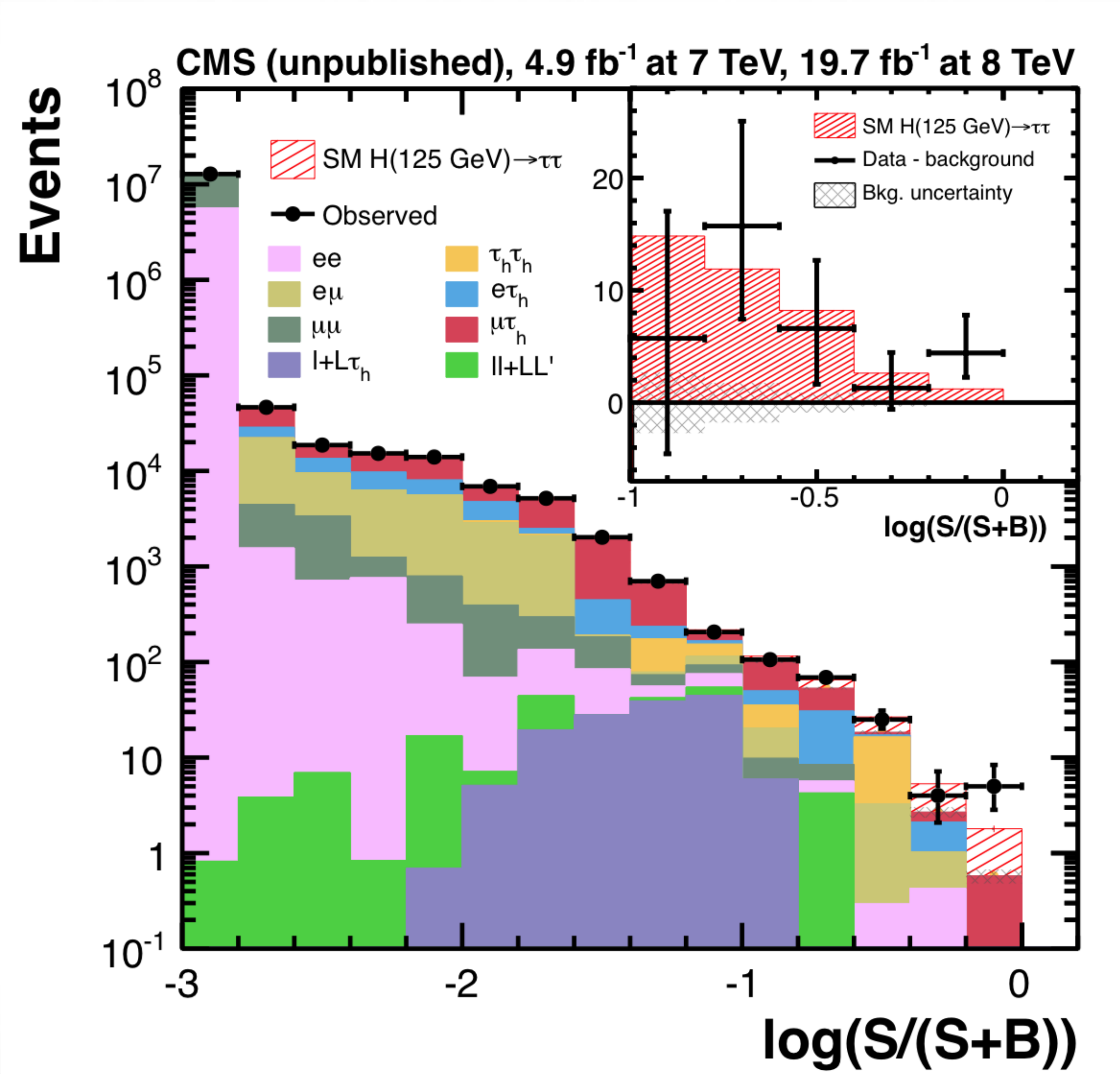
Systematics

on one slide

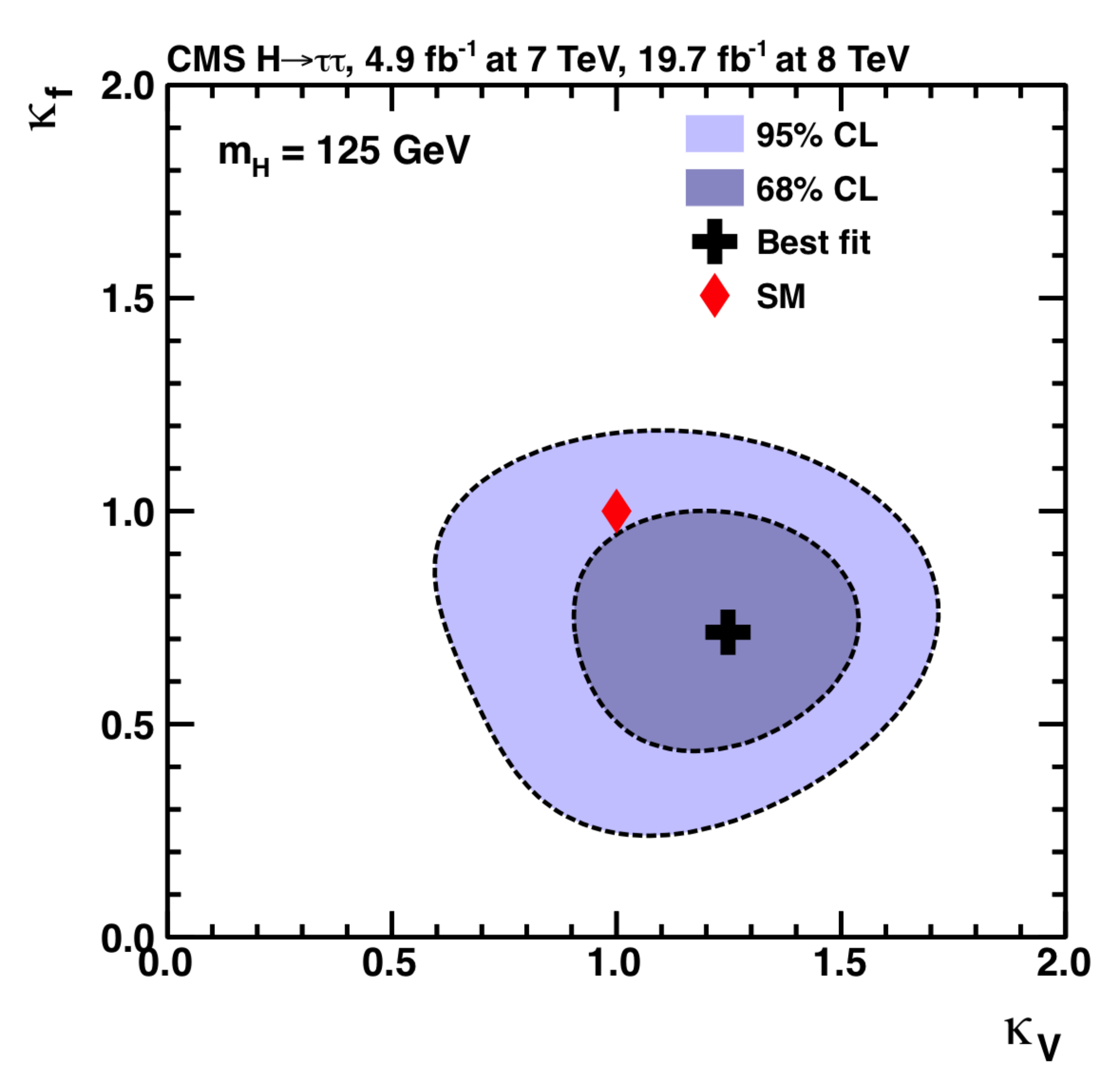


(J. Benitez)

S/(S+B) plot with all analysis bins



cV-cF



H → WW treated as signal

Contraint in κ_V comes from eμ VBF